INTRODUCCIÓN AL WATERMARKING

Elena Martínez Villacampa, Elisa Sayrol Clos

Departamento de Teoría de la Señal y Comunicaciones
E. T. S. Ingeniería de Telecomunicación de Barcelona

e-mail {elmarvi21@yahoo.com , elisa@gps.tsc.upc.es}

ABSTRACT

En los últimos años han aumentado los estudios sobre las técnicas denominadas de *watermarking* de contenidos multimedia. Con este artículo, se pretende dar una visión global del concepto de *watermarking*, explicar las técnicas empleadas para su uso y las principales utilidades que posee. Así mismo, se describe una estructura general del proceso de inserción del *watermark*, o marca de agua, en los datos originales y de la extracción o detección de la marca a partir de una imagen marcada. Pretendemos dar una visión completa y resumida de conceptos más utilizados y destacar algunas de las aplicaciones en las que se emplea como son la protección de copyright.

1. INTRODUCCIÓN

En la actualidad el uso cada vez mayor de Internet como vehículo para la distribución de datos multimedia y aplicaciones, así como los servicios on-line y las aplicaciones de comercio electrónico han provocado un aumento en el empleo de los datos digitales. Y aunque estos datos tienen numerosas ventajas frente a los analógicos, estos últimos son más seguros frente a reproducciones no autorizadas ya que sucesivas copias degradan notablemente el contenido original.

Ejemplos de información digital que no tiene protección ante posibles fraudes incluyen imágenes fijas, secuencias de vídeo, señales de audio, documentos de texto y gráficos.

Las técnicas de watermarking insertan una señal adicional, conocida como watermark o marca de agua, directamente en de los datos originales. Esta marca no es más que un mensaje idealmente imperceptible y de difícil extracción por parte de un usuario no autorizado. La información de este mensaje puede representar, por ejemplo, una secuencia binaria que contiene un número de serie, un logo, un número de tarjeta de crédito, una imagen o una firma.

Existen dos tipos de marcaje conocidos como marcaje frágil y marcaje robusto.

- Marcaje frágil: es una inserción débil de la marca, pensada para desaparecer tan pronto como modificamos demasiado el objeto a marcar. Este tipo de marcaje se emplea, por ejemplo, para la autentificación de datos.

- Marcaje robusto: este tipo de marcaje es imposible de eliminar sin degradar el objeto a niveles por debajo de los tolerables. Se emplea, por ejemplo, para la protección de copyright.

En la tabla 1 aparecen las propiedades más destacadas y deseables que han de tener los watermarks frágiles y robustos.

Marcaje frágil	Marcaje robusto
El watermark es alterado por la aplicación de las técnicas más comunes de procesado de imágenes.	El watermark permanece inalterado en una imagen, aún después de la aplicación de las técnicas más comunes de procesado de imágenes.
Es difícil para una persona no autorizada insertar un watermark falso.	El watermark es difícil de detectar por una persona no autorizada.
El watermark ha de ser rápidamente extraible por una persona autorizada.	El watermark ha de ser rápidamente extraible por una persona autorizada.
El watermark extraído indica donde han tenido lugar las alteraciones.	El watermark se ha de poder extraer, aún después de que la imagen marcada sufra alteraciones.

Tabla1. ***Propiedades principales del marcaje frágil y robusto si la marca es invisible [1]***

Existen también las denominadas técnicas semi-frágiles cuyas características consisten en ser robustas frente a transformaciones deseables, por ejemplo a la compresión, y frágiles a las manipulaciones mal intencionadas [11].

2. APLICACIONES DEL WATERMARKING

El watermarking tiene múltiples utilidades entre las que se pueden encontrar: [2, 3]

- Protección de Copyright (identificación del propietario): Para la protección de la propiedad intelectual, el propietario de los datos puede introducir una marca de agua representando la información de copyright de esos datos. Esta marca de agua puede probar su propiedad cuando alguien ha infringido sus copyrights. Actualmente la industria discográfica realiza grandes esfuerzos en introducir copyrights fiables en grabaciones en formato digital [12].

- **"Fingerprinting" (toma de impresiones digitales – identificación del comprador)**: El propietario puede emplear una técnica de 'fingerprinting' para rastrear la procedencia de copias ilegales. En este caso, el propietario de los datos digitales puede introducir diferentes marcas de agua en las copias de los datos que son suministrados a diversos clientes. Esta técnica es comparable a introducir un número de serie en los datos, que está relacionado con la identidad del cliente. Esto permite al dueño de la propiedad intelectual identificar clientes que han roto su contrato (licencia) suministrando estos datos a terceras partes.

- **Protección a la copia**: La información almacenada en una marca de agua puede controlar directamente dispositivos de grabación digitales con el fin de evitar copias ilegales. En este caso, el watermark representa un bit de prohibición de copias y los detectores de watermark en las grabadoras determinan si esos datos pueden ser almacenados o no. Actualmente la tecnología DVD incluye técnicas de watermarking en su sistema de protección de copias que contiene posibles mensajes como: "copia no permitida", "copiar una sola vez", "agotado el número de copias", ...[13]

- **Vigilancia de la transmisión**: Por ejemplo, para introducir marcas de agua en anuncios comerciales, un sistema de vigilancia automatizado puede verificar si los anuncios son transmitidos como han sido contratados.

- **Autentificación de datos**: Las marcas de agua frágiles pueden ser empleadas para comprobar la autenticidad de los datos. Este tipo de marcas indican si los datos han sido alterados y suministra información de localización como, por ejemplo, donde han sido alterados estos datos.

Las técnicas de watermarking se pueden emplear en otras aplicaciones además de como protección. Entre estas otras aplicaciones destacan:

- **Seguridad médica**: Puede ser una medida de seguridad útil incrustar la fecha y los datos del paciente en las imágenes médicas.

- **Esteganografía (ocultación de los datos)**: Las técnicas de watermarking pueden ser empleadas para la transmisión de mensajes privados y secretos.

Dentro de las técnicas comentadas previamente, vamos a enfocar este escrito al caso del empleo de técnicas de watermarking como sistema de copyright.

Las marcas empleadas deben ser robustas a cualquier alteración que se produzca en su contenido e incluir información suficiente para autentificar al propietario de estos datos. De este modo, se evita que personas no autorizadas se adjudiquen la propiedad de los datos a tratar o que manipulen esta información, eliminándola o haciéndola ilegible.

Las marcas que se introducen para autentificar al propietario legal de los datos a marcar, son los denominados watermarks robustos, que han de ser invisibles y soportar cualquier manipulación que se produzca en su contenido.

En la Fig. 1 aparece la imagen original 'lena', la imagen marcada y la marca introducida, que no es más que la diferencia entre la imagen original y la imagen marcada. Podemos ver cómo en la fig.1 b), la imagen marcada, no se aprecia diferencias significativas respecto a la fig. 1 a), imagen original. La fig. 1 c) representa la marca introducida. Para marcar esta imagen se han utilizado las técnicas desarrolladas en [16].

Figura. 1. a) *Imagen original 'lena',* ***b)*** *imagen marcada y* ***c)*** *marca introducida en la imagen original (diferencia entre la imagen original y la imagen marcada)*

Estos datos marcados serán posteriormente extraídos mediante un proceso de detección que se puede utilizar para demostrar la propiedad o no de los mismos.

3. PROPIEDADES DEL WATERMARKING EN IMÁGENES DIGITALES

En imágenes digitales, hemos de considerar algunos factores importantes a la hora de introducir una marca de agua. Es deseable un compromiso entre diversas características como son:

- Robustez.
- Invisibilidad.
- Cantidad de información.
- Seguridad.
- Recuperabilidad.

- Rapidez de extracción.
- Exclusión de ambigüedad.

· **Robustez**

Esta característica se refiere a la habilidad de la marca para permanecer en la imagen independientemente de que la calidad haya sido degradada, intencionada o no intencionadamente. Como ejemplos de degradaciones no intencionadas, podemos destacar, las aplicaciones relacionadas con el almacenamiento o la transmisión de datos, donde las técnicas de compresión son aplicadas a los datos para reducir las tasas de bits e incrementar la eficiencia. También, como degradaciones no intencionadas, encontramos las provocadas por conversiones analógico-digital, digital-analógicas, filtrados, muestreos, ... Otras manipulaciones incluirían las conocidas como distorsiones geométricas, que incluyen rotaciones, "cropping" (recorte de la imagen), deformaciones y desplazamientos.

Las características que ha de tener un sistema para que sea robusto son:

· La detección de la marca debería exigir el conocimiento de un secreto.
· Múltiples marcas no deben interferirse entre ellas.
· Las marcas deberían sobrevivir a cualquier posible ataque que no degrade la calidad a percibir.

· **Invisibilidad**

El algoritmo de watermarking debe introducir la marca de agua en la imagen de tal modo que ésta no afecte a la calidad de los datos que se encuentran debajo de ella. El proceso de introducción de la marca es realmente imperceptible si no somos capaces de distinguir los datos originales frente a los datos con la marca insertada.

Existe un compromiso entre la robustez y la invisibilidad. Cuanta más fuerza le demos a una marca, más visible se vuelve. Para ello se estudian máscaras perceptuales que, teniendo en cuenta el sistema visual humano (HVS), modulan el mensaje a insertar para utilizar la máxima potencia posible con la restricción de la imperceptibilidad. Por otro lado se estudian métodos de evaluación que nos den una medida de calidad visual a posteriori y que sean más apropiados que la relación señal a ruido (considerando como señal la imagen y como ruido la marca).

· **Cantidad de información**

Para lograr introducir la máxima información, ésta se reparte por toda la imagen. En algunas aplicaciones puede ser un factor determinante. Existen estudios teóricos sobre la capacidad máxima que se puede incluir en una imagen [4].

· **Seguridad**

En las técnicas de watermarking la marca a introducir ha de ser secreta y sólo accesible para las partes autorizadas.

Esta marca ha de ser resistente a posibles manipulaciones de usuarios no autorizados. Las personas que no posean derechos legales para su uso, no han de ser capaces de detectar ni descifrar la información introducida en ella.

· **Recuperabilidad**

En sistemas de detección ciega, donde no disponemos de la imagen original, la marca introducida ha de tener unas propiedades concretas que permitan su posterior recuperación por los usuarios autorizados que poseen permiso para su empleo.

Además, también hemos de tener en cuenta que la probabilidad de error en detección ha de ser baja para evitar que se produzcan falsas detecciones positivas.

· **Rapidez de extracción**

Es deseable que el proceso de detección sea rápido para toda persona autorizada o para el propietario original de las imágenes digitales. Sobre todo, en los sistemas que trabajan en tiempo real, es un factor determinante.

· **Exclusión de ambigüedad**

Para aplicaciones de copyright y fingerprinting, la marca introducida ha de identificar con total seguridad a la persona o empresa propietaria legal de los derechos de la propiedad intelectual de la imagen o al comprador.

4. MODELADO DE UN SISTEMA DE WATERMARKING COMO UN SISTEMA DE COMUNICACIONES

Un sistema de watermarking se puede interpretar como un sistema de comunicaciones convencional, en el que para el proceso de marcaje tenemos un canal de comunicación de datos (el contenido mismo de la imagen) que nos sirve como soporte de la información a introducir (marca de agua). En el proceso de detección, la imagen marcada es la señal recibida, mientras que la marca es la señal transmitida. Además, como en los sistemas ciegos la imagen original no es conocida en recepción, se puede considerar ésta como un ruido aditivo que dificultará la detección de los datos. En la fig. 2 podemos ver un diagrama de bloques que representa un sistema de watermarking con detector ciego (en el que no disponemos de la imagen original) representado como un modelo de comunicacione

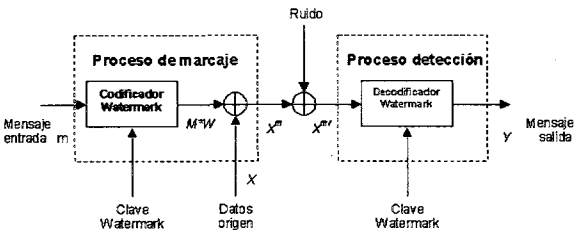

Figura2. Sistema de watermarking con detector ciego representado como un modelo de comunicaciones

El conocimiento del canal en el proceso de marcaje puede utilizarse para intuir y mejorar el proceso de detección. En este caso el modelo es un canal de comunicaciones con información lateral [3].

4. 1 Proceso de marcaje

La marca de agua se introduce de forma aditiva en la imagen, (aunque algunos autores utilizan modelos multiplicativos [17]). Así, en general, la señal marcadaserá:

$$X^m = X + a \cdot M \cdot W$$

donde X^m es el vector que representa los datos marcados, X representa el vector de los datos originales, W es la marca de agua que se introduce en los datos a tratar, M representa la máscara perceptual, que modifica los datos a marcar según sus características y, finalmente, a es un valor que pondera la marca introducida para controlar la potencia de la marca en los datos.

El proceso de marcaje puede introducir la marca directamente en los datos originales o bien en una versión transformada de los mismos. Entre los distintos dominios empleados podemos destacar, el temporal (para secuencias de audio) o espacial (para imágenes); y otros dominios transformados como son la transformada discreta de Fourier, DFT, la transformada discreta del coseno, DCT, [5, 6] o la transformada discreta *wavelets*, DWT [7, 5].

Las técnicas en dominios transformados son muy utilizadas debido a su estructura natural para incorporar conocimientos perceptuales en los algoritmos de marcaje y porque muchas técnicas de compresión como JPEG, MPEG2 (DCT) o JPEG-2000 (DWT) trabajan en esta misma estructura, lo que permitirá recuperar la marca con una decodificación parcial y no total de los datos. El uso de estos dominios hará posible obtener unos mejores resultados en sistemas donde se produce compresión.

La marca introducida puede ser un valor binario o real. En algunas aplicaciones, a esta marca se le aplica una clave secreta. Sin el conocimiento de ella, resultará difícil extraer o alterar el mensaje insertado sin destruir los datos originales. Los sistemas que emplean técnicas de espectro ensanchado se usan frecuentemente para aplicaciones de este tipo [8]. Habrá otras, en cambio, donde no será deseable el empleo de claves. Los sistemas QIM (Quantization Index Modulation) serán los aconsejados en estos casos. [9, 10]

4. 2 Proceso de detección

La detección o verificación de la marca recibida se refiere al proceso de realizar una decisión binaria en el detector. Comprobar si una marca concreta se encuentra o no en los datos recibidos.

Existen diferentes esquemas de detección de la marca. Los que incluyen clave secreta y los datos originales, clave secreta pero sin el original (detección ciega) o los que no poseen ni clave ni datos de origen.

Para recuperar la información introducida en los datos marcados, se puede emplear la detección por máxima verosimilitud. También haremos referencia a un caso particular, la detección por correlación dada su simplicidad y buenos resultados [11]. Este método es óptimo en los casos con presencia de ruido blanco Gaussiano aditivo. La correlación normalizada es muy similar a la anterior. En este caso, previo a la realización del producto promedio entre los elementos de dos vectores, estos son normalizados. De este modo, disponemos de un detector más robusto a posibles cambios en los datos marcados.

En el detector, los valores de detección se calculan para cada una de las marcas que forman el mensaje introducido en la imagen. El mensaje más probable será aquel cuyas marcas tengan los valores de detección mayores. Si el valor obtenido es menor a un umbral, el detector no reconocerá la existencia de la marca de agua, mientras que si el valor es mayor a ese mismo umbral se considerará como una detección positiva de la watermark.

Con este método cabe la posibilidad de obtener falsos positivos, se detecte una marca cuando no existe, o de falsos negativos, que no detecte una que existe. Para muchas aplicaciones es más importante tener pocos o ningún falso positivo, aún a costa de aumentar el número de falsos negativos. Por ese motivo el valor umbral escogido, se puede modificar en relación a la probabilidad de detección y la probabilidad de falsos positivos.

Por otro lado, el desarrollo y evolución de los algoritmos de detección se ha beneficiado del estudio de los ataques que se producen en los esquemas de watermarking y en las posibles soluciones (contraataques) que se crean [14, 15]. Un ejemplo de ataque es la pérdida de sincronización que provoca que el cálculo de la correlación no sea efectivo. El contraataque propuesto consiste en usar partes de la marca para introducir una sincronización conocida en los datos a tratar.

5. CONCLUSIONES

El uso cada vez más extendido de Internet ha implicado una revolución en el manejo de información y datos digitales, hecho que ha propiciado el empleo de técnicas de *watermarking*, para asegurar fiabilidad y seguridad en los datos.

Las técnicas de watermarking se emplean para múltiples aplicaciones entre las que podemos destacar la protección de copyright, la protección a la copia o la autentificación de datos. Estas técnicas consisten en la introducción de un mensaje oculto en los datos a tratar. Muchas de ellas se apoyan en los conceptos de sistemas de comunicaciones y del sistema perceptual humano.

El watermarking proporciona robustez, frente a posibles ataques, a todas aquellas aplicaciones que lo emplean. Existirá un compromiso entre cantidad de información insertada en los datos a tratar y robustez. También existirá un compromiso entre robustez y percepción visual.

Existen aún muchas líneas de investigación abiertas tanto por lo que respecta a las actuales limitaciones que presenta el proceso de marcaje y como el de detección. El número de posibles ataques que se pueden producir en los datos marcados, como son las distorsiones geométricas, son incalculables. Todos estos factores crean un campo de investigación en desarrollo para conseguir un aumento de la efectividad de las técnicas de watermarking.

6. REFERENCIAS

[1] F. C. Mintzer, G. W. Braudaway y M. M. Yeung, "Effective and Ineffective digital watermarks". IEEE in Proceedings, the International Conference on Image Processing, Oct. 1997, vol. 3, pp. 9-12.

[2] Gerhard C. Langelaar, Iwan Setyaman, y Reginald L. Langendijk, "Watermarking Digital Image and Video Data". IEEE Signal Processing Magazine, Sept.2000, pp. 20-46.

[3] Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom, "Digital Watermarking". Morgan Kaufmann. 2002.

[4] Pierre Moulin, M. Kivanç Mihçak, "A Framework for Evaluating the Data-Hiding Capacity of Image Sources". IEEE Transactions on Image Processing, Sep. 2002, vol.2, pp. 1029 – 1042.

[5] International Workshop on Information Hiding, 1996.

[6] E. Koch, J. Zhao, "Towards robust and hidden image copyright labeling". Nonlinear Signal Processing Workshop, Thessaloniki, Greece, 1995.

[7] M. Swanson, B. Zhu, A. Tewfik, "Multiresolution scene-based video watermarking using perceptual models". IEEE J. Select. Areas Commun., May. 1998, vol. 16, pp. 525-539.

[8] I. J. Cox, J. Kilian, T. Leighton, T. Shamoon, "Secure spread spectrum watermarking for multimedia". NEC Research Institute, Princeton, NJ, Technical Report 95-10, 1995.

[9] B. Chen, G. Wornell, "Dither modulation: A new approach to digital watermarking and information embedding". SPIE, Security and Watermarking of Multimedia Contents, San Jose, CA, Ene. 1999, vol. 3657, pp. 342-353.

[10] B. Chen, G. Wornell, "Preprocessed and postprocessed quantization index modulation methods for digital watermarking". SPIE, Security and Watermarking of Multimedia Contents II, San Jose, CA, 2000, vol. 3971, pp. 48-59.

[11] Christine I. Podilchuk, Edward J. Delp, "Digital Watermarking: Algorithms and Applications". IEEE Signal Processing Magazine, Jul. 2001, pp. 33-46.

[12] L. Boney, A. Tewfik, K. Hamdy, "Digital watermarks for audio signals". IEEE Proc. Multimedia, 1996, pp. 473-480.

[13] J. Bloom, I. Cox, T. Kalker, J. P. Linnartz, M. Miller, C. Traw, "Copy protection for dvd video". Proc. IEEE, nº 87, Jul. 1999, pp. 12667-12760.

[14] S. Craver, N. Memon, B.-L. Yeo, M. Yeung, "Can invisible watermarks resolve rightfulownerships?". SPIE Electronic Imaging '97: Storage and Retrieval of Image and Video Databases". 1997, pp. 310-323.

[15] I. Cox, J.-P. Linnartz, "Some general methods for tampering with watermarks". IEEE J. Select. Areas Commun., May. 1998, vol. 16, pp. 587-593.

[16] Elena Martínez, "Watermarking de Imágenes en Color en el dominio wavelet". Proyecto Fin de Carrera. ETSETB. UPC. 2003.

[17] M. Barni, F. Bartolini, A. de Rosa, A. Piva, "A New Decoder for the Optimum Recovery of Nonadditive Watermarks". IEEE Transactions on Image Processing, May. 2001, vol. 10, nº 5, pp. 755-766.

AUTORES

Elena Martínez Villacampa *nació en Barcelona el 21 de octubre de 1976. Finalizó los estudios de Ingeniería Técnica de Telecomunicaciones, especialidad en Sistemas de Telecomunicación en la Escuela Universitaria Politécnica del Baix Llobregat (UPC) en 1999. En la actualidad está realizando el proyecto de fin de carrera sobre* watermarking de imágenes en color en el dominio wavelet *en el departamento de Teoría de Señal y Comunicaciones de la Escuela Técnica Superior de Ingeniería de Telecomunicaciones de Barcelona (UPC) y trabaja como becaria en la empresa Vodafone.*

Elisa Sayrol Clols *recibió el título de Ingeniera de Telecomunicación y el título de Doctora de la Universitat Politècnica de Catalunya en los años 1989 y 1994 respectivamente. Desde 1990 hasta 1993 hizo cursos de postgrado en la* Northeastern University, *y fué investigadora visitante en la* University of Southern California, *donde trabajó con el Profesor C.L. Nikias.Actualmente es profesora titular de la Universitat Politècnica de Catalunya, en la Escola Superior d'Enginyeria de Telecomunicació de Barcelona e imparte el curso* Senyals i Sistemes 1, *así como el curso de doctorado* Processat Digital d'Imatge. *Sus intereses investigadores incluyen estadísticas robustas i de orden superior, el análisis de imagen y video y las técnicas robustas de* watermarking.